

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/982,828

1600

DATE: 03/05/2002
TIME: 06:29:48

INPUT SET: S36790.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

RECEIVED

MAR -5 2002

TECH CENTER 1600/2900

ENTERED

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Thurber, Denise
Zeng, Bin

(ii) TITLE OF INVENTION: Coding Sequences of the Human
BRCA1 Gene

(iii) NUMBER OF SEQUENCES: 72

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Morgan Lewis & Bockius LLP
(B) STREET: 1111 Pennsylvania Avenue N. W.
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: USA
(F) ZIP: 20004

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/982,828
(B) FILING DATE: 2001-10-22
(C) CLASSIFICATION:

(vii) PRIORITY APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/074,453
(B) FILING DATE: 1998-05-06

(A) APPLICATION NUMBER: US 08/798,691
(B) FILING DATE: 1997-02-12

(A) APPLICATION NUMBER: US 08/598,591
(B) FILING DATE: 1996-02-12

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/982,828DATE: 03/05/2002
TIME: 06:29:49

INPUT SET: S36790.raw

47
48 (viii) ATTORNEY/AGENT INFORMATION:
49 (A) NAME: Michael S. Tuscan
50 (B) REGISTRATION NUMBER: 43,210
51 (C) REFERENCE/DOCKET NUMBER: 44921-5053-01-US
52

53 (ix) TELECOMMUNICATION INFORMATION:
54 (A) TELEPHONE: 202-739-3000
55 (B) TELEFAX: 202-739-3001
56
57
58

59 (2) INFORMATION FOR SEQ ID NO:1:
60

61 (i) SEQUENCE CHARACTERISTICS:
62 (A) LENGTH: 5711 base pairs
63 (B) TYPE: nucleic acid
64 (C) STRANDEDNESS: Not Relevant
65 (D) TOPOLOGY: linear
66

67 (ii) MOLECULE TYPE: cDNA
68

69 (vi) ORIGINAL SOURCE:
70 (A) ORGANISM: Homo sapiens
71 (B) STRAIN: BRCA1 (omil)
72

73 (viii) POSITION IN GENOME:
74 (A) CHROMOSOME/SEGMENT: 17
75 (B) MAP POSITION: 17q21
76
77
78
79

80 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
81

82	AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCCTCAGA TAACTGGGCC	60
83		
84	CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA	120
85		
86	TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA	180
87		
88	TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC	240
89		
90	ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT	300
91		
92	GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC	360
93		
94	AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTCGGAGT	420
95		
96	ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG	480
97		
98	AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG	540
99		

RECEIVED
MAR -5 2002
TECH CENTER 1600/2900

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/982,828DATE: 03/05/2002
TIME: 06:29:49

INPUT SET: S36790.raw

100	AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA	600
101		
102	CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG	660
103		
104	AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG	720
105		
106	ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG	780
107		
108	CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC	840
109		
110	CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT	900
111		
112	ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA	960
113		
114	GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA	1020
115		
116	AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT	1080
117		
118	GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG	1140
119		
120	ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT	1200
121		
122	CAGAGAATCC TAGAGATACT GAAGATGTTT CTTGGATAAC ACTAAATAGC AGCATTGAGA	1260
123		
124	AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG	1320
125		
126	GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG	1380
127		
128	AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA	1440
129		
130	TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT	1500
131		
132	TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAATT AAGCCATGTA ACTGAAAATC	1560
133		
134	TAATTATAGG AGCATTGTGT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA	1620
135		
136	AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG	1680
137		
138	CAGATTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC	1740
139		
140	AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT	1800
141		
142	CTATTGAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA	1860
143		
144	AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACCTGAA TTAAATATCC	1920
145		
146	ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC	1980
147		
148	ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA	2040
149		
150	TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAA GTACAACCAA ATGCCAGTCA	2100
151		
152	GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA	2160

Line	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6	Distance
153							
154	GTAACAAGCC	AAATGAACAG	ACAAGTAAAA	GACATGACAG	TGATACTTTC	CCAGAGCTGA	2220
155							
156	AGTTAACAAA	TGCACCTGGT	TCTTTTACTA	AGTGTTCAAA	TACCAGTGAA	CTTAAAGAAT	2280
157							
158	TTGTCAATCC	TAGCCTTCCA	AGAGAAGAAA	AAGAAGAGAA	ACTAGAAACA	GTAAAGTGT	2340
159							
160	CTAATAATGC	TGAAGACCCC	AAAGATCTCA	TGTTAAGTGG	AGAAAGGGTT	TTGCAAACTG	2400
161							
162	AAAGATCTGT	AGAGAGTAGC	AGTATTTTAC	TGGTACCTGG	TACTGATTAT	GGCACTCAGG	2460
163							
164	AAAGTATCTC	GTTACTGGAA	GTTAGCACTC	TAGGGAAGGC	AAAAACAGAA	CCAAATAAAT	2520
165							
166	GTGTGAGTCA	GTGTGCAGCA	TTTGAAAACC	CCAAGGGACT	AATTCATGGT	TGTTCCAAAG	2580
167							
168	ATAATAGAAA	TGACACAGAA	GGCTTTAAGT	ATCCATTGGG	ACATGAAGTT	AACCACAGTC	2640
169							
170	GGGAAACAAG	CATAGAAATG	GAAGAAAGTG	AACTTGATGC	TCAGTATTTG	CAGAATACAT	2700
171							
172	TCAAGGTTTC	AAAGCGCCAG	TCATTTGCTC	TGTTTTCAAA	TCCAGGAAAT	GCAGAAGAGG	2760
173							
174	AATGTGCAAC	ATTCTCTGCC	CACTCTGGGT	CCTTAAAGAA	ACAAAGTCCA	AAAGTCACTT	2820
175							
176	TTGAATGTGA	ACAAAAGGAA	GAAAATCAAG	GAAAGAATGA	GTCTAATATC	AAGCCTGTAC	2880
177							
178	AGACAGTTAA	TATCACTGCA	GGCTTTCTCTG	TGGTTGGTCA	GAAAGATAAG	CCAGTTGATA	2940
179							
180	ATGCCAAATG	TAGTATCAAA	GGAGGCTCTA	GGTTTTGTCT	ATCATCTCAG	TTCAGAGGCA	3000
181							
182	ACGAAACTGG	ACTCATTACT	CCAAATAAAC	ATGGACTTTT	ACAAAACCCA	TATCGTATAC	3060
183							
184	CACCACTTTT	TCCCATCAAG	TCATTTGTTA	AACTAAATG	TAAGAAAAAT	CTGCTAGAGG	3120
185							
186	AAAACTTTGA	GGAACATTCA	ATGTCACCTG	AAAGAGAAAT	GGGAAATGAG	AACATTCCAA	3180
187							
188	GTACAGTGAG	CACAATTAGC	CGTAATAACA	TTAGAGAAAA	TGTTTTTAAA	GGAGCCAGCT	3240
189							
190	CAAGCAATAT	TAATGAAGTA	GGTTCCAGTA	CTAATGAAGT	GGGCTCCAGT	ATTAATGAAA	3300
191							
192	TAGGTTCCAG	TGATGAAAAC	ATTCAAGCAG	AACTAGGTAG	AAACAGAGGG	CCAAAATTGA	3360
193							
194	ATGCTATGCT	TAGATTAGGG	GTTTTGCAAC	CTGAGGTCTA	TAAACAAAGT	CTTCCTGGAA	3420
195							
196	GTAATTGTAA	GCATCCTGAA	ATAAAAAAGC	AAGAATATGA	AGAAGTAGTT	CAGACTGTTA	3480
197							
198	ATACAGATTT	CTCTCCATAT	CTGATTTTCA	ATAACTTAGA	ACAGCCTATG	GGAAGTAGTC	3540
199							
200	ATGCATCTCA	GGTTTGTTCT	GAGACACCTG	ATGACCTGTT	AGATGATGGT	GAAATAAAGG	3600
201							
202	AAGATACTAG	TTTTGCTGAA	AATGACATTA	AGGAAAGTTC	TGCTGTTTTT	AGCAAAAGCG	3660
203							
204	TCCAGAGAGG	AGAGCTTAGC	AGGAGTCCTA	GCCCTTTCAC	CCATACACAT	TTGGCTCAGG	3720
205							

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/982,828DATE: 03/05/2002
TIME: 06:29:50

INPUT SET: S36790.raw

206	GTTACCGAAG	AGGGGCCAAG	AAATTAGAGT	CCTCAGAAGA	GAAGTTATCT	AGTGAGGATG	3780
207							
208	AAGAGCTTCC	CTGCTTCCAA	CACTTGTTAT	TTGGTAAAGT	AAACAATATA	CCTTCTCAGT	3840
209							
210	CTACTAGGCA	TAGCACCGTT	GCTACCGAGT	GTCTGTCTAA	GAACACAGAG	GAGAATTTAT	3900
211							
212	TATCATTGAA	GAATAGCTTA	AATGACTGCA	GTAACCAGGT	AATATTGGCA	AAGGCATCTC	3960
213							
214	AGGAACATCA	CCTTAGTGAG	GAAACAAAAT	GTTCTGCTAG	CTTGTTTTCT	TCACAGTGCA	4020
215							
216	GTGAATTGGA	AGACTTGACT	GCAAATACAA	ACACCCAGGA	TCCTTTCTTG	ATTGGTTCTT	4080
217							
218	CCAAACAAAT	GAGGCATCAG	TCTGAAAGCC	AGGGAGTTGG	TCTGAGTGAC	AAGGAATTGG	4140
219							
220	TTTCAGATGA	TGAAGAAAGA	GGAACGGGCT	TGGAAGAAAA	TAATCAAGAA	GAGCAAAGCA	4200
221							
222	TGGATTCAAA	CTTAGGTGAA	GCAGCATCTG	GGTGTGAGAG	TGAAACAAGC	GTCTCTGAAG	4260
223							
224	ACTGCTCAGG	GCTATCCTCT	CAGAGTGACA	TTTTAACCAC	TCAGCAGAGG	GATACCATGC	4320
225							
226	AACATAACCT	GATAAAGCTC	CAGCAGGAAA	TGGCTGAACT	AGAAGCTGTG	TTAGAACAGC	4380
227							
228	ATGGGAGCCA	GCCTTCTAAC	AGCTACCCTT	CCATCATAAG	TGACTCCTCT	GCCCTTGAGG	4440
229							
230	ACCTGCGAAA	TCCAGAACAA	AGCACATCAG	AAAAGCAGT	ATTAAGTTCA	CAGAAAAGTA	4500
231							
232	GTGAATACCC	TATAAGCCAG	AATCCAGAAG	GCCTTTCTGC	TGACAAGTTT	GAGGTGTCTG	4560
233							
234	CAGATAGTTC	TACCAGTAAA	AATAAAGAAC	CAGGAGTGGA	AAGGTCATCC	CCTTCTAAAT	4620
235							
236	GCCCATCATT	AGATGATAGG	TGGTACATGC	ACAGTTGCTC	TGGGAGTCTT	CAGAATAGAA	4680
237							
238	ACTACCCATC	TCAAGAGGAG	CTCATTAAGG	TTGTTGATGT	GGAGGAGCAA	CAGCTGGAAG	4740
239							
240	AGTCTGGGCC	ACACGATTTG	ACGGAAACAT	CTTACTTGCC	AAGGCAAGAT	CTAGAGGGAA	4800
241							
242	CCCCTTACCT	GGAATCTGGA	ATCAGCCTCT	TCTCTGATGA	CCCTGAATCT	GATCCTTCTG	4860
243							
244	AAGACAGAGC	CCCAGAGTCA	GCTCGTGTG	GCAACATACC	ATCTTCAACC	TCTGCATTGA	4920
245							
246	AAGTTCCCCA	ATTGAAAGTT	GCAGAATCTG	CCCAGGTCC	AGCTGCTGCT	CATACTACTG	4980
247							
248	ATACTGCTGG	GTATAATGCA	ATGGAAGAAA	GTGTGAGCAG	GGAGAAGCCA	GAATTGACAG	5040
249							
250	CTTCAACAGA	AAGGGTCAAC	AAAAGAATGT	CCATGGTGGT	GTCTGGCCTG	ACCCAGAAG	5100
251							
252	AATTTATGCT	CGTGTACAAG	TTTGCCAGAA	AACACCACAT	CACTTTAACT	AATCTAATTA	5160
253							
254	CTGAAGAGAC	TACTCATGTT	GTTATGAAAA	CAGATGCTGA	GTTTGTGTGT	GAACGGACAC	5220
255							
256	TGAAATATTT	TCTAGGAATT	GCGGGAGGAA	AATGGGTAGT	TAGCTATTTT	TGGGTGACCC	5280
257							
258	AGTCTATTAA	AGAAAGAAAA	ATGCTGAATG	AGCATGATTT	TGAAGTCAGA	GGAGATGTGG	5340

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/982,828

DATE: 03/05/2002
TIME: 06:29:50

INPUT SET: S36790.raw

Line	Error	Original Text
------	-------	---------------